CHAPTER 2

SCIENTIFIC METHOD : A SOLUTION

Discovery . . . is a byproduct of making things simpler.[1]

Introduction

When analysing any body of information larger than the average mind can encompass in a single span, it is probably best to group the evidence in some simple sort of way, so that each of its aspects can initially be assessed separately, much as the pieces of an unknown jigsaw puzzle are best grouped first into those of like colour, form, etc. In this way, the amount of information in each group can be reduced to more manageable proportions, and if the groups are severally sufficient to encompass all of the information, the overall problem itself is then reduced to the lesser one of finding how the answers to each group can be fitted together. To use Koestler's analogy, even the patient monkey at the typewriter would have a better chance of composing his Shakespearian sonnet if the keys on the typewriter contained whole words rather than single letters. [2]

I should like to start by proposing a method along these lines for synthesising diagnoses in clinical medicine, since this is the point from which my own views first began to evolve.

Synthesis of Clinical diagnoses

Pattern recognition methods, though extensively used in clinical diagnosis, have two serious. First, unless we are dealing with classical disease patterns, there is the danger of pigeon-holing patients within too-restricted categories. And since each individual has the 'irritating' habit of being different from the next, this frequently leads to an oversimplified view of the disease condition in particular patients. Of course, experience will teach us a number of disease variants, and perhaps with it the ability to recognise when something new lies beyond their scope, but it will not help much in determining what that new condition actually is. This is a particular problem for the academic clinician, for he tends to be referred only very unusual cases, and yet he must always be in a position to make some sensible comment upon them, and to recognise in them what is new. As well, the pattern-recognition approach is an exceedingly difficult and cumbersome way to learn medicine as a student, and indeed to impart it as a teacher. Given the vast patterns of disease and the rapidily-expanding progress of medical knowledge, an approach of this type would now almost require a life-time's experience before anyone could be thought of as a really competent diagnostician!

These points struck home particularly when I first came to the University of Tasmania, for I did so during a time of changeover to a new undergraduate curriculum. There, we were faced with introducing our medical students to clinical diagnosis a year earlier than before (in the third year of their then 6-year course), and to avoid completely bewildering them, we needed to impart at least some general guide-lines for clinical problem-solving. After contemplating this difficulty and reading around the subject for some time, I found I had to reject many of the suggested approaches, including the now popular problem-based learning[3] in favour of what I believed to give the first real clue to the solution, and one which came from the much earlier writings of Gowers. [4] He held that, without necessarily knowing the names of particular conditions, one could begin to understand

something about them by considering quite separately the anatomical site of the lesion from its pathological nature. This general approach seemed potentially very valuable, for not only did it allow an initial separation of clinical information into two different groups, but the groups themselves also fitted well with our students' preclinical background in the basic disciplines of anatomy and general pathology. Using such a method we might begin to make these basic subjects work better for us in reaching a diagnosis, rather than just using them in an explanatory sort of way after the event, as so often is the case. What Gowers recommended was to:

"Discard in the first instance all attempts to identify or to name, and try instead to read the malady, tracing the symptoms to the seat of their cause, and discerning the nature of the morbid process by their character and course."

As a clinical neurologist, Gowers had a powerful influence on his field, and its way of teaching, and others since have used and extended his views. [5] But apart from some very perceptive general analyses [6] such modifications have on the whole been disappointing, especially from the viewpoint of *individual* patient diagnosis: the principle for that has never been more clearly stated than by Gowers himself.

In extending Gowers' teachings to current-day clinical diagnosis, I saw it necessary to broaden the number of his categories to allow the inclusion of the greater breadth of clinical information accumulated since that time, and also to take account of the increased number of preclinical disciplines into which that information might now be classified. At the same time, I had to be careful not to expand the number of categories too far, since this would risk losing the very simplicity of Gowers' approach. The problem, then, was to develop several simple basic medical science categories into which clinical data might conveniently be grouped, each relatively independent of the other yet, taken as a whole, broad enough to encompass the full potential range of that data. Here, Kipling[7] had as much to say to me as he has, in different contexts, to others such as Richard Asher:[8]

"1 keep six honest serving men
(They taught me all I knew);
Their names are What and Why and When
And How and Where and Who." [7]

Using Gowers' principle of separating clinical information into groups, and Kipling's advice on economy in their number and simplicity in their form, I arrived at the following four questions which may be formulated to help solve clinical diagnostic problems. They are as follows:

1. *Where is the lesion?* (Anatomical diagnosis).

2. *What is the pathological nature of the condition?*
(Pathological diagnosis).

3. *How does it affect the various bodily systems?*
(Functional diagnosis).

4. *Why did this particular patient get it?*
(Aetiological diagnosis).

These four broad categories are simple, relatively independent of each other, and at the same time severally sufficient to encompass most clinical information. They have the added virtue of being based in the separate disciplines of Anatomy, General Pathology, Physiology and preclinical disciplines relevant to Aetiology including the topics of biochemistry, microbiology, special pathology and epidemiology.

The first three categories of diagnosis are used to decide from the clinical information what, in gene terms, the condition actually, is. The fourth relates to why the patient has such a condition in the first place,

includes such possible relevant predisposing factors as, for example, “risk factors” in patients with heart att (myocardial infarction). We shall see how the method may be made to work in a moment, but I first want to stress the importance of keeping to a minimum not only the number of broad categories, but also the quest asked within each of them. Neurologists, for example, [5] have separately developed Gowers’ views and, ind have put forward the same diagnostic categories I delineate here. However, they surround them with such a bewildering and unordered array of associated questions, that there is a grave risk of losing the essence of Gowers’ original approach. I think Kipling[7] would agree, for later in the same verse at the end of “The Elephant’s Child” he goes on to say:[7]

> “But different folk have different views;
> I know a person small-
> She keeps ten million serving-men,
> Who get no rest at all!
> She sends ‘em abroad on her own affairs,
> From the second she opens her eyes-
> One million Hows, two million Wheres,
> And seven million Whys!”

This brings me to my second major point. It is true that one must often ask many questions to obtain enough clinical information within each of the four broad categories to reach a diagnosis. However, some questions are more important than others, and it is possible to structure questioning in an ordered, and even hierarchic, way — the first questions always being broader, and the later ones focusing down on the finer levels of detail. Such an approach to the general process of learning is also alluded to by Bruner [1] in his dissertation “On Knowing: Essays for the Left Hand”, where he suggests that much can be learned from games like “Twenty Questions”. Those who succeed best there tend initially to ask broad questions that allow the *general* area of the solution to be defined. Only then do they go on to secondary questions about specific aspects of detail. There would be no point, for example, in asking whether any unknown object was a kitchen sink, if its general description were to be found classified under the broader headings of Animal / Living!

I believe that one sees examples of the use and success of this general approach in the everyday workings of many men of practice, particularly the self-taught mechanic who, despite his lack of formal knowledge (or perhaps because of it), comes to recognise that his complex machines consist essentially of broad categories of functions in much the same way as I have outlined, and as Pirsig has discussed in detail. [9] Self-taught or otherwise, the good mechanic is rapidly able to localise any fault by first determining which *system is* involved. Moreover he does this efficiently by testing the one component in each system whose functioning depends on the integrity of all the others within it. An enquiry aimed first at this level gives the vital initial information as to whether there is *anything* wrong with that particular system. Once the faulty system has been localised by this process, a similarly ordered enquiry can again be conducted within it by aiming tests at successively lower and lower levels of its functionally hierarchic organisation. For example, when a old car is broken down, testing whether or not there is spark across the spark-plug gap when the motor is turned over tells, with few exceptions, whether there is *any* fault within the ignition system. If the spark is present and active, there is generally no need to waste time checking other components of that system such as the battery, coil, condenser, distributor, spark plug leads etc., because the problem lies elsewhere. On the other hand, if there is no spark across the gap the fault will have been localised to the ignition system and the region of that fault can then be systematically analysed. The same can be said of the fuel system, which is the other major resource component of a motor engine. This, too, can be assessed most efficiently by checking it at the “end of the line,” as it were (in this case the carburettor). Again this first gives essential information about whether there is *anything* wrong at all with the system. I am convinced that electricians, electronic engineers and many other tradesmen work in a very similar way, indeed to such an extent that these principles often become second nature or intuitive, and hence unfortunately lost to easy analysis and explanation. When an electrician takes an appliance which is reportedly not working and first tests it in an alternative power supply, he is clearly illustrating the simple principle of first localising the system at fault. Of course, as

my own machines and motors, to my chagrin, continue to remind me, one has to admit that much more is involved than this in practice. Second, some sites in some systems are more conveniently, if less appropriately, tested initially than others nearer the “end of the line”, because they are more accessible. Also, with experience, certain patterns of aberrant performance will point strongly to some particular fault without the need to employ any special method of tracing it Nonetheless, I am convinced that the general approach outlined can and does help many men of practice, especially in the analysis of unusual or difficult problems.

Anatomical Diagnosis

This method of localising a fault is very relevant to the delineation of an anatomical diagnosis in clinical medicine, i.e. in answering the question, “Where is the lesion?” Again the first questions should be those which allow an assessment of which organ system is at fault, before we proceed to look at finer levels of detail. For example, we may very well draw false conclusions about a patient presenting with the symptom of weakness if we assume that the problem lies within the central nervous system when it is actually a muscle disorder. Of course, errors of this sort are usually rectified and nothing may be lost except time, inconvenience, and cost of needless investigation. But in an emergency situation, such as a suddenly unconscious patient, much more is at stake. There, the patient’s life may depend on our rapidly localising the fault, and in my view the best approach is not very different from that of our bush mechanic. Thus, in this context, palpating the carotid pulse would provide an excellent indication of whether there was anything seriously wrong with the general circulatory system of blood supply to the brain. If the pulse were absent, we could immediately localise the fault to the cardiovascular system; if present we could go straight on to other systems, thus avoiding time lost in needlessly checking the various components within it. On the other hand, if the pulse were present and the patient cyanosed (blue), we would have localised the trouble to the respiratory system. In that case, respiration itself being broadly divided into those functional components related to ventilation on the one hand and those related to gas exchange with the blood on the other, we could then set about further localising the site of the problem within the respiratory system by testing, in turn, these and lower orders of its hierarchic structure and function in a similar way. Other systems relevant to adequate brain performance could also be tested quickly by analysing each at their point of output on which the brain function depends. Only in this way can the fault be localised rapidly enough to avoid potential disaster. And how often are we, as doctors trained in the detail of medicine, embarrassed by the first-aid man at the scene of an accident who localises the trouble whilst we are still pondering its details!

Of course, there are principles other than this that must be used in any such situation — the importance of looking for reversible causes is one, but another I emphasise concerns with the efficient localisation of any fault in answering the question “Where is the lesion?”

To determine an anatomical diagnosis by this general method, I suggest the following (hierarchic) order of questioning:

1. Which of the bodily systems is/are involved? If there are more than one, the next question must be asked.

2. Is this a multi-system disease, or can the whole picture be explained by disease in a single system with secondary manifestations elsewhere?

Here we see the first signs of an important interaction between our four broad diagnostic categories — a point we shall discuss in some detail below — because to answer this question we may have to look first at whether some parts of the clinical information may be more related to the effects of the condition than to its primary cause (e.g. the secondary manifestations of chronic liver impairment in cirrhosis of the liver). Solving such a problem will frequently be helped by parallel consideration of the Functional diagnosis along with the Anatomical one. In cases of special difficulty, two simple rules often help. One is that what came first in time is more likely to be related to cause than to effect. The other is the principle of Occam’s razor (“It is vain to do with more what can be

done with less") that would tend to favour a single system disease with secondary manifestations over a multi-system disorder in cases of any doubt.

This parallel consideration of the Anatomical and Functional diagnoses should not cause any confusion, since they often have to be considered together in any case, particularly in dealing with remote organs like the brain, where the only guide to the anatomical site of the lesion may be the functional aberrations it causes.

3. Where in the system is the fault? Once the system, (e.g., cardiovascular) and the organ within it (e.g. the heart) has been localised, then more detailed questions can be asked about the region involved — always remembering that such questioning should follow the descending order of functional and/or structural organisation within that system — so that the Anatomical diagnosis will gradually become more and more precise. Even in the most obscure of conditions, a statement can always be made about Anatomical diagnosis at some level using this method.

The clinical information relevant to making an Anatomical diagnosis comes from features in both the history and the clinical examination. The contribution of each will depend on the system involved, but broadly speaking the historical features that contribute most are the major symptoms themselves, their quality, and the factors which aggravate or relieve them. Depending again on the system, the clinical examination may give either direct evidence about the site of the lesion, as in the abdomen where any enlarged organ may be felt with the hand, or indirect evidence as in the respiratory and cardiovascular systems, through, say, auscultation. Yet other information may come from a degree of overlap between the Anatomical diagnosis and each of the other three broad diagnostic categories, particularly Functional diagnosis. This latter aspect will be considered in more detail below.

Pathological Diagnosis

This general hierarchic approach to questioning is also valid for each of the other three broad diagnostic categories. Thus, in answering the question "What is the general pathological nature of the lesion?" Gowers[4] recognised that some questions were initially more important than others, and the one to which he gave particular emphasis was the time-intensity relationships of the presenting complaint. Through this, we can build up an important picture (even a mind's-eye graph) of the onset, duration, intensity, variability, and progression of the condition. The onset of an illness is especially relevant to understanding general clinical pathology. Symptoms continuing over many years, and progressing only very slowly (i.e. very chronic or long term), most often signify a degenerative condition, though it may also be hereditary, or due to chronic toxins, poisons, or chronic deficiency states. Those whose onset is over a period of many months (chronic) rather than many years, and which follow a steadily progressive course, by and large have either a chronic inflammatory (inflammation = '-itis', as in bronchitis) or a chronic neoplastic (cancerous) basis. In contradistinction to the degenerative disorders both of these latter conditions tend to be associated with weight loss. On the other hand a distinction may be made between them by looking for the presence of fever and other indications of the inflammatory state. Conditions of more rapid onset may be subdivided on the basis of whether they came on over a few weeks to a few days (acute), or a few seconds (hyperacute). If over a few seconds, the condition is likely to be due to a rupture or obstruction of a hollow tube somewhere within the body — be it vascular, visceral or otherwise — and once such a hyperacute onset has been established, further questions can be asked to allow a choice between these alternatives. For example, if the symptoms are transient and clear up within a few minutes, it is more likely that the event is due to an obstruction than to the less-reversible condition of rupture.

After determining whether the condition is acute, sub-acute, chronic, relapsing, episodic, periodic, etc., and thereby minimising the range of pathological entities which might be concerned, the next phase of our enquiry in this category should be aimed at delineating which is the most likely. For example, if a condition can be broadly defined as being 'acute/progressive', we must then go to the next level of hierarchic dissection and ask whether this is a process of (acute) inflammation, toxic poisoning, ischaemic necrosis (injury/death of tissue from lack of blood supply), hypersensitivity reaction, etc., and in choosing between these possibilities we must again ask broad questions first.

For example we should ask: “Are there *any* of the hallmarks of inflammation?” before assuming the condition to be ‘bacterial’ in nature. There is no point at all in even considering that unless the general and/or local indicators of inflammation are present e.g. fever, local swelling, pain, redness and tenderness. Even then, our hierarchical approach to questioning should continue. Thus, having established a condition to be acute/inflammatory, we would want to know whether there was an increase in the circulating neutrophil polymorphonuclear leucocytes, and whether immature forms and toxic granulation changes were present among them, before concluding that the condition was likely *bacterial*. So the hierarchy of our general pathological diagnosis in this case so far would be, acute/inflammatory/bacterial, and further dissection could follow to allow the type of organism involved to be defined.

There are other aspects of the clinical information which bear on defining the general pathological nature of the condition, including the texture, regularity, and ease of delineation of any palpable organ (e.g. the ‘rock-hard’, irregular liver of secondary cancer) as well as information which comes from any degree of overlap of the Pathological diagnosis with the other three broad diagnostic categories (Anatomical, Functional, & Aetiologica! diagnoses). Indeed our discussion of the case diagnosis in the previous paragraph gives an example of how two of these — the Pathological (acute/inflammatory) and the Aetiological(bacterial?pneumococcal) diagnoses in this case — may merge as we descend to the finer levels of detail. Overall, though, the time-intensity relationships, and the presence or absence of weight loss and fever yield by far and away the most useful information in delineating the broad outline of the general pathology involved within this category.

Functional Diagnosis

Our third category, the functional diagnosis, is the final one suggested for finding what, in general terms, the condition confronting us actually is. Specifically, we ask,

“How does this condition affect the various bodily functions?”

And as we have already seen, we may well need to address this question early in difficult cases and ask, in a slightly different way, “Which of the clinical features may relate to causation, and which can be seen as possible functional or secondary effects?” For example, high blood pressure associated with impairment of renal function can be viewed either as its cause or its effect, and the information will have to be looked at both ways around initially (“Which came first in time?”) before going any further down the line on the basis of tacit assumptions.

The Functional diagnosis contributes to the overall diagnosis in several ways. First, as we have just seen, it may help our analysis of the clinical information by separating features that can be related to cause from those more likely due to effects. Second, the Functional diagnosis often overlaps importantly with the Anatomical diagnosis. This is particularly so in systems where direct evidence on anatomy is hard to come by (e.g. through direct manual palpation) during physical examination. In such cases, the knowledge that there are clinical features suggesting impairment of a particular organ system helps to confirm suspicions on other grounds that this is indeed the system anatomically involved. Thus, in a patient presenting with non-specific central chest pain consistent with the involvement of virtually any of the mediastinal structures, the finding of heart failure would point strongly to the cardiovascular system as the likely anatomical site of the problem.

Some systems are so remote from direct access to examination that we become almost entirely reliant on the presence of functional changes to make any Anatomical diagnosis at all. The classical case of this type is the brain, where it almost entirely through our knowledge of neuro-physiology and the way neuronal changes affect organ function elsewhere — e.g. the musculo-skeletal system — that gives us our neuro-anatomical diagnosis, and moreover often does so with an accuracy of within a few millimetres. Finally, as a diagnostic category in its own right, the functional diagnosis may bear strongly on our approach to the patient’s treatment, for even where the underlying disease is not amenable to therapy its functional effects may be (e.g. the value of oxygen therapy in patients with emphysema).

The first three broad diagnostic categories are essential for an adequate assessment of the presenting condition. Our approach to obtaining clinical information within each category will vary from case to case, but one normally tries to get some idea of the organ system involved (either directly, or more indirectly through assessment of its effects) from the initial history, and an idea of

the general Pathological nature of the condition. From the physical examination, one expects to find further information about the Anatomical and Functional diagnoses, and as a result, one may then need to go back and question the patient further. By the end of that time, however, we should be in a position to make some sensible comment within each of these three diagnostic categories so as to synthesise a general diagnosis about the presenting complaint As another example, a diagnosis might now read, “Hepatic/acute inflammation/secondary liver failure.”

Aetiological Diagnosis

Having gained some idea of the general diagnosis, we next have to delve into the background and ask,

“Why did this particular patient get this condition?”

This is very helpful in determining the root cause of the problem, for example high blood pressure and /or cigarette smoking as background “risk” factors predisposing to acute myocardial infarction

I have stressed in this chapter that we should not have too many broad categories of diagnosis, but this consideration of background aetiological factors suggests that we might usefully expand our number of “honest serving men” to Kipling’s six, by including a further two questions, particularly if we can do so under already existing categories (in this case the Aetiological), as follows:

“**WHEN**?” By this I mean,

“Why did this particular patient get this condition *when* he did?

This question is especially useful in unearthing any factors which might have *precipitated* an *acute* clinical episode, particularly in acute-on-chronic conditions such as myocardial ischaemia/infarction, where any considerations about background aetiology are usually confined to the more long-term atheroma “risk factors” (high blood pressure, high plasma cholesterol, etc.) to the exclusion of what could possibly have precipitated such a dramatic-onset clinical event as myocardial infarction. And if no precipitating factors become evident to us, we should ask our patient what he/she thinks might have provoked the condition. Of course, the answers can often be difficult to interpret, but they are important for us to know, and are sometimes revealing and thought-provoking.

“**WHO**”? By this I mean WHO has the disease?

And here I am thinking of the words of Francis Scott Smythe:

“To know what kind of person has a disease is as essential as to know what kind of a disease a person has.”

This is an important question, not only because the personality and character strengths and weaknesses of patients will be of the utmost importance in determining their response to any disease process - and, with that, their ability to cope - but because an individual patient’s reaction to life events, its stresses and stressors, may be crucial as an initiating factor (perhaps more important than we realise) actually precipitating his/her clinical condition For example, acute episodes of bronchial asthma are undoubtedly precipitated in some patients by psychological stress. The interests of time and space do not allow us to pursue this matter further here, but it is important[1], and is something I will take up in later chapters when we come back to look more broadly at mechanisms in disease.

As far as the structure of our enquiry is concerned, clinical information relevant to the Aetiological diagnosis may not be forthcoming until after the initial history-taking and physical examination have been completed, i.e. it will not be until we have gained a general idea of what the condition actually is that we will be in a position to delve more deeply into its background cause.

Diagnostic Category Overlap

[1] “What is spoken of as a ‘clinical picture’ is not just a photograph of a man sick in bed; it is an impressionistic painting of the patient surrounded by his home, his work, his relations his friends, his joys, sorrows, hopes and fears.” — F.W.Peabody, 1881-1927.

This is another important to diagnosing in categories. Thus, the Aetiological category of diagnosis can be very helpful in allowing us to question whether the very setting in which the general diagnosis has been made is reasonable. For example, no matter how strongly the clinical information may point to a general diagnosis of acute myocardial infarction, one would hesitate before applying that label to a 15 year old boy with no obvious predisposing "risk factors". The pain in our young patient may be arising from the cardiovascular system right enough, and the acuteness of onset of the episode may be well-established, but is it perhaps the pericardium rather than the myocardium which is involved, so does he really have acute pericarditis?

Actually, every one of the four broad diagnostic categories may interact so as to question or strengthen each other. We have already seen how the Functional diagnosis may interact strongly with the Anatomical one, particularly in organs remote from the direct reach of our examining senses. Also, as Gowers[4] has pointed out, there are some combinations of Anatomical and Pathological diagnoses which do not fit well together. Thus, damage to peripheral nerves by toxic poisoning would not fit well with a nerve lesion occurring in, say, one limb. On the other side of the coin, a knowledge about pathology will sometimes favour one particular Anatomical diagnosis over another, such as in the predilection for cerebral haemorrhage to occur in the internal capsule of the brain rather than in the anatomical site from which it may be difficult to distinguish clinically, namely the pre-frontal cortex. If, therefore, we obtain a history of rapid onset of relentlessly progressive cerebral symptoms suggesting (cerebral) haemorrhage as the Pathological diagnosis, then an internal capsular Anatomical site for the lesion is generally more likely.

Another example: in a patient presenting with sudden onset dysphasia and weakness in the right lower face and upper limb, now improving, we can say that this is a hyperacute event (the 'What?' of diagnosis) involving the left prefrontal cortex (the 'Where?'). Moreover, given such sudden onset, it is likely due to a blockage or rupture of a hollow tube somewhere within the system. And, since it is improving, blockage is more likely than rupture. Now, if we were dealing with the abdomen, where there are many hollow tubes, this would not help much. But in the brain the only important hollow tubes are vascular ones. This means that we can now look at the anatomical diagnosis not just in neuroanatomical terms, but in terms of neurovascular anatomy. From that, it is a simple step for us to realise that the territory involved is the left middle cerebral artery, where that we are dealing with some sort of vascular obstructive episode, either local thrombosis or embolic from elsewhere. In this way, we have been able to greatly sharpen the anatomical diagnosis using the pathological one.

Once these four diagnostic categories have been satisfactorily answered, it remains to fit them together. Often this process is simple and straight-forward, e.g. acute(Path) myocardial(Anat) infarction(Path) with secondary ventricular dysrhythmia and left heart failure(Funct), on a long-term background background of severe hypertension(Aetiol). But on occasions, as we shall see, the separate diagnostic categories themselves, and the way they seem to fit best together, can cause some surprises and even lead to unanticipated conclusions. This process may be akin, albeit on a somewhat lesser level, to Koestler's view that the uncovering of novel points of intersections between different frameworks of reference to any body of information may well lie at the basis of the "Eureka act" in science, laughter in humour, and elevation to a new plane of aesthetic understanding in the creative arts. [2] I do not suggest that this technique will always give the correct diagnosis in particular patients, or that a new insight will necessarily be obtained into disease processes in general. Nor will it always be necessary in making diagnoses, particularly where the pattern of the disease is clearly recognisable. But in learning the diagnostic process as a student and in dealing with difficult diagnostic problems as a physician, it can be extremely helpful. Importantly, it can enable us to at least '*corner* ' the diagnosis i.e. to narrow it down to *some* level of refinement within each of the four diagnostic categories, so as to allow a better approach to the problem overall.

All of this must seem rather like being given a written description of how to tie a neck-tie, where an illustration would far better demonstrate the point. Let me therefore now give some examples of what I mean.

CASE DISCUSSIONS

Case 1

Let us first consider the problem of a 12 year old boy who presents to hospital with a one week history of increasing malaise, loss of appetite, nausea and occasional shivers and sweats, in association with increasing shortness of breath on exertion, swelling of the ankles, and the passage of a reduced quantity of urine, described by his mother as being of 'smoky' appearance. He had previously been well, except for a sore throat a week or so before the onset of this illness. On physical examination, he is noted to have some slight puffiness of the eyes, swelling of the ankles, and a raised temperature of 38 C. His blood pressure is moderately elevated at 150/100 mm. Hg; there is a raised jugular venous pressure, and some moist sounds are heard at the lung bases. Abdominal examination reveals no abnormality. Urine volume is small and contains protein, red cells, including red cell casts and some neutrophil polymorphonuclear leucocytes. No bacteria are seen in the urine. Central nervous system examination normal, including retinal vessels in the optic fundi.

A pattern-recognition approach to diagnosis would see this as a classical case of acute post-streptoccocal glomerulonephritis. But what of our young student who knows no names, and whose understanding can only come through his knowledge of anatomy, general pathology, and physiology, etc. He may of course stumble across the diagnosis if he recognises its pattern, but if he does not, let him ask the following questions in turn

1. Where is the lesion? (Anatomical Diagnosis)

This is a condition affecting multiple systems in the body, including the kidneys (urinary abnormalities), the lungs, (basal moist sounds), the circulation (raised systemic arterial and venous pressures), and the face and ankles (puffiness and swelling). Because of this, our method tells us to first consider whether this is truly a multi-system disease or whether it is primarily an affliction of a single organ system with secondary manifestations elsewhere. The principle of Occam's razor would make us lean initially in the second direction, and we should examine each of the symptoms and signs in turn to see whether they fit best as cause or effect. We might first consider for example that the high blood pressure is the primary abnormality with involvement of the kidney from secondary effects. It would help in this respect to know from the patient's history whether the urinary abnormalities began before or after the onset of the high blood pressure, but in this case it seems unlikely that so moderate and elevation of blood pressure would cause such dramatic renal effects; this alone would be sufficient to throw doubt on the association being this way around. On the other hand, some primary abnormality of the kidney might account for the elevated blood pressure and some of the other aspects of the patient's condition as well. In examining such possible interrelationship, we should recall the importance of considering the Functional diagnosis side by side with the Anatomical one, so that rather than analyse the Pathological diagnosis next, we should now go on to consider the Functional diagnosis.

2. How does the condition affect the Function of the various bodily systems? (Functional diagnosis).

In this case the real question is "Can we fit the findings together satisfactorily on the basis of single organ disease with secondary manifestations ?". So far we have not been able to do this on the basis of the hypertension being primary, but with primary renal disease many of the features could be seen as secondary manifestations of impaired kidney function. Thus, the modest elevation of blood pressure might be explained by renal disease involving the juxta-glomerular apparatus, which controls blood pressure through renin release. The swelling of the face and hands and the raised jugular venous pressure are consistent with a general increase in extra-cellular fluid volume, and this could be explained by a decreased renal glomerular filtration. Moreover, a reduction in glomerular filtration rate could also account for the reduced urine volume. The proteinuria, too, could be caused by an impaired integrity of the glomerular capillary membrane, which normally holds protein back from filtration into the urine. Alternatively, the oedema, raised venous pressure, moist sounds in the lung and proteinuria might all be considered as secondary manifestations of primary cardiac failure. But there is no abnormality detectable in the heart, and the only factor which we know

might cause the heart to fail, namely the high blood pressure, is again only very moderately raised. Also, it would be difficult for primary cardiac failure to explain all of the symptoms, particularly the red blood cell casts and white cells in the urine. To invoke primary cardiac disease as a way of seeing therefore, would give us only a partial explanation of the clinical picture and, as a result, we would again be faced with violating the principle of Occam's razor by having to make two separate diagnoses. Primary renal disease with secondary effects from impaired renal function could, on the other hand, explain all of the information and is therefore the preferred way of seeing. Having made this preliminary analysis on the basis of alterations in organ function, we are now in a position to go back and look more closely at the Anatomical diagnosis. In this respect most of the secondary or functional effects discussed so far have been related to aberrations of glomerular function, so if our student's basic approach is right, the Functional diagnosis has helped to localise the Anatomical site of the lesion to the glomerulus.

So far, then, our diagnosis is of a primary condition of the kidney involving the glomerulus and producing secondary impairment of glomerular function.

3. What is the Pathological nature of the lesion?

Firstly, it was acute in onset, and secondly there is evidence of inflammation, though the nature and cause of this is not yet clear. It might be a bacterial inflammation, or one secondary to ischaemic necrosis, toxic damage or immunological change, and one of the pieces of information we need to make this decision — not clinical but still vital — is whether there is a high circulating blood polymorphonuclear leucocyte count, with toxic neutrophil granulation which might indicate a bacterial cause. Locally in the urine, inflammatory cells are present, but although this too would be consistent with an acute local inflammatory process, there are no bacteria to support its having a direct bacterial cause. If available, examination of the urine for immunological abnormalities might prove helpful.

So far therefore we can say that this is an acute inflammatory condition of the kidney involving the glomerulus and producing a secondary impairment of glomerular function. Now we have established this, we can think a little more about causation.

4. Aetiological diagnosis: Why did the patient get this condition?

Without knowing any patterns of disease, this may be difficult, but because by our method any events occurring in the background to an illness must be considered in relation to causation, the patient's sore throat a week or two before the onset of this illness would have to be looked at very hard in this regard.

We have therefore reached a broadly similar diagnosis to that determined by the pattern-recognition approach. In some respects we have done even better, because we have only drawn conclusions that are reasonable on the basis of the evidence in *this* patient. Thus, the diagnosis of 'glomerulonephritis' suggests that elements other than the glomeruli are involved in the inflammatory process, and there is no evidence for that in this case. Also, our analytical approach has left open those areas where there is still room for reasonable doubt. Thus, the Anatomico-Pathological diagnosis "acute glomerul-itis" is warranted, but the question of its underlying causation (Aetiology) is left unresolved. It may indeed be post-streptococcal, but until further evidence is available, that is an unwarranted assumption. Note also that in contradistinction to the pattern-recognition approach, our diagnosis comments on the secondary Functional effects of the condition; these may be highly relevant to treatment as well as diagnosis.

We have therefore, through this approach, 'cornered' the diagnosis as far as is reasonable, and clearly indicated areas where we need more information. This seems to me better than the general hanging of labels which don't necessarily fit the *individual* patient well. Moreover, this analytical or synthetic approach is particularly useful in obscure cases, and essential in those patients presenting with conditions the like of which we have never seen before or which are new and not yet even described.

Case 2

In the case history just discussed, I deliberately picked a somewhat difficult example as a severe test of the method. A more straight-forward one, but one which illustrates the dangers of hidden assumptions implicit in the pattern recognition approach, is as follows.

A 38 year old woman presents to hospital with a history of having developed an ‘influenza-like’ illness four days beforehand. However, instead of improving, this gradually worsened and she began to develop increasing malaise, lethargy, shortness of breath, shivers and sweats, and eventually a sharp pain made worse by breathing over the front of the right lower chest. She had also developed a cough productive of thick yellowish-brown phlegm. Physical examination reveals a temperature of 39°C, rapid ‘grunting’ respiration, pallor, some confusion, and evidence of central cyanosis (blue colour). There is a recent ‘cold sore’ on the upper lip. Examination of the respiratory system shows a pleural friction rub over the right lower chest anteriorly, in association with signs of consolidation in the underlying lung tissue.

Like the first example, this case history is exceedingly brief to say the least, but it will suffice to make the point.

According to the pattern-recognition approach, this is classical pneumococcal lobar pneumonia,almost recognisable from the end of the bedside, given the ‘herpes simplex’ cold sore on the lip, and the character of the respiration. But let us pause a moment and see what conclusions we can reach using our approach. The Anatomical site of the lesion is straight-forward. Examination of the respiratory system shows consolidation in the middle lobe of the right lung; the patient also has pleuritic chest pain with an associated pleural friction rub (causing the ‘grunting’ respiration), so that the condition involves the overlying pleura as well. In general Pathological terms, the illness has been of acute onset and there is evidence for an inflammatory nature, both general (shivers, sweats, and fever) and local (purulent and probably blood stained sputum). The Functional diagnosis includes secondary cyanosis (oxygen lack or hypoxaemia) and some resulting confusion. The diagnosis so far is therefore of an acute inflammatory condition involving the middle lobe of the right lung and the overlying pleura, with the secondary manifestations of hypoxia and confusion. There is an associated cold sore on the lip, probably due to the herpes simplex virus.

Our general conclusions are therefore similar to those reached by the pattern-recognition approach. The conclusions may not be as brief, but in some respects they are better. On the one hand our diagnosis goes further where the information allows. For example it gives us more specific information about the Anatomical site of the lesion (right middle lobe, pleura) and about its secondary Functional effects, both of which may bear importantly on the patient’s treatment. In other important respects it does not go as far. We have not, for example, assumed that this condition is bacterial, let alone pneumococcal. Our approach has emphasised that the Pathological diagnosis is less than complete, and in need of more information on the nature of the inflammatory process. This would include an examination of the blood white cell count for the general hallmarks of bacterial inflammation, as well as a local investigation of the sputum (by microscopy etc.) for any pathogenic organisms. Finally through its functional category, our diagnosis focuses attention on the important secondary aspects of illness, for in this instance the patient urgently needs oxygen therapy. We can say little about the underlying cause or aetiology as yet, but our approach serves to remind us that, when the patient’s condition improves, we must delve further into the background in this respect, not only looking at the question “Why?” but also “When?” and “Who?”

There is no doubt that pattern-recognition does, with experience, play an important part in clinical diagnosis, and where it gives the correct answer I have no quibble at all with it, because it certainly can be an efficient process. But rather like an off-the-rack suit, it does not necessarily fit well in all places, and without further analysis, it might be hard to tell wherein these places lie. At the very least, therefore, we should pause after making such a diagnosis and reflect on what each of the words and syllables in any diagnostic label mean and how well they fit the patient under consideration

When we think about how these labels read, they usually encompass most of our four broad diagnostic categories in any case, though unfortunately in fixed combination. In some ways, therefore, all we may sometimes have to do is examine the diagnostic label itself and see how well each of its components fits the individual patient. For example, the label “acute viral hepatitis” encompasses three of our categories; Pathology (acute), Aetiology (viral), Anatomical (hepar), Pathology (inflammation = ‘itis’), although it is striking that, as with so many diagnostic labels, the

Functional category is completely lacking. Even in cases that appear classical, it can be instructive to put each of the diagnostic syllables under scrutiny. In doing so, one may occasionally find that one particular aspect is a surprisingly poor fit, perhaps overlooked in the ordinary way because of the commonness of the general condition and the classical features exhibited by the patient in other respects. Such an approach might cause us to replace the word ‘viral’ in the above diagnosis, for example, by the word ‘alcoholic’.

Even where the pattern of the disease is absolutely classical, the label applied to it does not always give full insight. Labels are certainly convenient, useful as a guide to prognosis, and, valuable in prescribing general treatment, but unfortunately once applied, they tend to inhibit all further thought about the nature of the condition in the *individual*.

Diagnostic categories and disease mechanisms

I now move away from the area of individual patient diagnoses to the broader area with which much of this book will be concerned, namely the analysis of data in medical science as a means of understanding more about the general nature of human disease. Fortunately, as we shall see, our four broad diagnostic categories will continue to be helpful in this respect. Sometimes their application merely adjusts the perspective of our thinking without making any really fundamental change. Such is the case with the diagnostic label “chronic duodenal ulcer” for example. The pain associated with this condition is characteristically periodic in nature, coming and going intermittently with periods of freedom from pain that may last several months. As a student I was taught to recognise this periodicity as a very important characteristic by which to diagnose this condition, and very helpful in differentiating it from other upper abdominal complaints. But if we stop to consider that, by our method, the time-intensity relationships of any symptom tell us a great deal about the general Pathological nature of the condition underlying it, we can immediately see that such a time-course does not suggest a chronic process at all, but an acute relapsing and remitting one. This suggests that we are dealing with a process of episodic healing and breakdown of the duodenal mucosa rather than a truly chronic inflammatory process. This may seem simple, almost self-evident, but the pattern-recognition approach to diagnosis certainly does not highlight the point. And in a research setting, any such inherently relapsing nature is often completely forgotten when investigating the underlying disease mechanism. For example, many of the ‘auto-immune’ diseases run a remitting and relapsing course, and one is therefore entitled to ask how self-reacting antibodies alone can cause disease which comes and goes in such manner. These and other more general points will be taken up in subsequent chapters. But to show that there really are some basic problems with our general concepts in some areas of medical science bearing close analysis in this respect, let us end with one further case example, and consider what our young medical student might make of the following patient.

Case 3

An 18 year old girl presents with a history of recurrent ‘fitting’ over the past three months. The attacks are sometimes heralded by a twitching in the left hand, but are then followed within a few seconds by sudden loss of consciousness, and, according to witnesses, spasm of the limb muscles with arching of the back succeeded by coarse regular jerking movements of all the limbs, and then urinary incontinence and tongue biting. The attacks usually last no longer than a minute, are often followed by drowsiness and headache, and occasionally by automatic behaviour not subsequently remembered by the patient. Usually the whole episode subsides within an hour and the patient returns completely to normal as confirmed by neurological examination in between attacks. The history from the parents is that the attacks often seem to come on during a period of stress, or following unusual tiredness. The patient has been well in the past and has had no previous illnesses,

operations, or head injury. There is a family history of epilepsy. The only drug she has been taking is oral contraceptive therapy, for the past two years.

As trained physicians, it is not difficult to recognise this as focal epilepsy rapidly spreading to become a complete grand mal fit. Our general approach to this along classical lines would be to look for an underlying cause, and if we find none, to label the condition as “idiopathic grand mal tonic/clonic epileptiform seizures” (The more obscure the condition the more we tend to borrow from other languages to describe it, and the longer the name itself).

But let us see how we might understand the problem by avoiding labels and attempting instead to construct a diagnosis through our four broad categories. Firstly, given that the attack starts by involving the left hand, the focus of the problem is probably in the motor representation of the hand within the right pre-frontal cortex. However, because the jerkiness spreads rapidly to involve both arms and legs and is associated with loss of consciousness, it would appear that the initially localised process soon becomes generalised throughout the central nervous system as a whole. Postulates about the pathological nature of the condition itself should be consistent with the very abrupt onset. Now this onset could, I suppose, represent some sort of mechanism not seen in other organs, and in that sense the term ‘idiopathic’ might be appropriate. But by applying Occam’s razor and working for the time being with known mechanisms, such an onset would lead us to suspect that the events were precipitated by some sort of obstruction or rupture of hollow tubes somewhere within the brain. In this particular organ, the only hollow tubes are vascular ones, and this suggests a sudden rupture or obstruction of a blood vessel as the cause. Moreover, since the attacks are rapidly reversible, vascular obstruction seems much more likely. Venous or lymphatic obstruction would hardly cause such an abrupt loss of cerebral function, so that we are left by exclusion with an acute arterial obstructive cause. Instead of calling the condition “idiopathic epilepsy” therefore, or describing it as a “spontaneous neuronal discharge”, if we stay within the framework of our approach we can see it as the result of an acute reversible vascular obstruction within the territory of blood supply of the right pre-frontal cortex (i.e. within the branches of the right middle cerebral artery).

Now, in contradistinction to our analysis of chronic duodenal ulcer, we are suddenly in the midst of controversy, for this is not at all what most physicians or medical researchers see as the basis of “idiopathic epilepsy” — despite some proponents of this view over the years, and despite the fact that epilepsy has many analogies with migraine, which is now generally accepted as having such a vascular basis. I will discuss the possible vascular cause of epilepsy later in this book, but for the moment I wish merely to make the point that this very simple approach has led us to a viewpoint quite different from the traditional, and this strikes me as being healthy, indeed necessary, in a condition whose origins remain obscure. And in a field such as this, I see no reason to postulate unknown mechanisms when known ones can give a perfectly adequate explanation for them. Of course, until it is shown that such attacks are indeed associated with, and preferably preceded by, an appropriate regional reduction in blood flow, one would not expect everyone to accept this as the true cause. But it is a much more useful heuristic than the term ‘idiopathic’. Of course, other mechanisms have been put forward to explain the neuronal instability, but I would point out that very few of these give any indication of why the condition should be so episodic in its time-course, why so sudden in onset, and why so localised anatomically to particular areas of the brain. It seems to me that an acute reversible vascular obstruction supplies a very adequate explanation of all of these, and until such an explanation is found wanting, it should not be lightly rejected for others, no matter how fashionable.

I will go on in the next chapter to consider the analysis and synthesis of data in the various fields of *medical science*, using our four diagnostic cornerstones as foundations for the approach, and always bearing in mind that finding different ways of seeing them in different combinations may be very helpful in giving us a new approach to understanding the disease in question. Since, at best, this method can only allow us to reach probable conclusions, we cannot hope to be always right, but if initially we go no further than the information allows, we should at least be able to ‘*corner* ’the beginnings of a solution. And even where we need to go beyond, for example in fitting the four broad diagnostic categories together, provided we make our point of departure from interpretation clear, any contentious overall synthesis should not call into question the sub-conclusions within each of the

categories themselves. Part II of this book will be concerned largely with the area of analysis and interpretation of existing medical evidence. In Part III, we will need to go beyond this, for I believe it will be constructive and creative to allow our four broad diagnostic categories to interact so that we can look at their possible combinations in different ways. Not only is this constructive, but it is at times necessary, for as Poincare has pointed out, “When a logician has resolved each demonstration into a host of elementary operations, all of them correct, he will not yet be in possession of the whole reality; that indefinable something that constitutes the unity will escape him completely.” [10] I am not advocating holism in the strict sense by quoting this view, but merely suggesting that the possession of the whole reality may be dependent on ways of seeing different combinations of each of its parts. This may be as when several areas of a jigsaw puzzle have been solved, yet the difficult problem of how best to put them together remains. In some ways, of course, this now takes us back to the realm of the undefined rules of hypothesis-formulation, for to put various categories together in the best way may indeed have to be a somewhat random hypothetico-deductive process. But as we will now have the preliminary solutions to each of its parts, the number of pieces we have to contend with in this way will be reduced to more manageable proportions, and even the random approach may now be creative. Perhaps, then, we will be more like Koestler’s monkey given whole paragraphs instead of words for each typewriter key.[2]

Finally, I apologise to those not totally familiar with all of the medical terms used in the various case discussions above, though perhaps in reality the acid test of the method proposed may be whether or not interested readers including postgraduate medical and science students have been able to better understand the conditions discussed as a result.

References

1. Bruner J. *On knowing. Essays for the left hand*. New York: Atheneum Press; 1976.
2. Koestler A. *The act of creation*. London: Hutchinson; 1964. p. bibl., illus.
3. Spalding WB. *Revilatizing medical education: Mcmaster medical school, the early years 1965-1974*. Philadelphia: B.C. Decker; 1991.
4. Gowers W. Clinical lectures on diseases of the nervous system. *Lancet*. 1892; **i**:9.
5. Tyrer JH, Eadie MJ. *The astute physician*. Amsterdam: Elsevier; 1976.
6. Feinstein AR. An analysis of diagnostic reasoning. I. The domains and disorders of clinical macrobiology. *Yale J Biol Med*. 1973; **46**:212-32.
7. Kipling R. The elephant' s child. In: *Just so stories*. London: MacMillan and Co, 1931.
8. Avery Jones F. *Richard asher talking sense*. Bath, U.K.; 1972.
9. Persig RM. *Zen and the art of motor cycle maintenance*. London: Corgi Books; 1976.
10. Poincare H. *Science and method*. London: Thomas Nelson and Sons; 1914.